Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
`http://siba-ese.unisalento.it/index.php/ejasa/index`
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v9n1p246

Dependence modeling using EVT and D-Vine

By Sesay, Liu

Published: 26 April 2016

Electronic Journal of Applied Statistical Analysis
Vol. 09, Issue 01, April 2016, 246-266
DOI: 10.1285/i20705948v9n1p246

Dependence modeling using EVT and D-Vine

Regina Baby Sesay*[a] and Baokon Liu[b]

[a] *Mathematics and Statistics Department, Njala university, Njala, Sierra Leone*
[b] *School of Statistics, Southwestern University Chengdu, China*

Published: 26 April 2016

Extreme occurrences such as extreme gains and extreme losses in financial market are unavoidable. Accurate knowledge of dependence on these extremes can help investors to adjust their portfolio mix accordingly. This paper therefore focuses on modeling dependence in extreme gains and losses in a portfolio consisting of three assets using D-Vine Copula. The concept of extreme value theory (EVT) is used to identify the sets of extreme gains and extreme losses in each asset contained in the portfolio. For the three assets considered, a total of six sets (3 sets of extreme gains and 3 sets of extreme losses) are use in the dependence modeling. The inference function for margin (IFM) approach is used in the analysis. The first stage of the IFM is the modeling of the marginal distributions; this is done using the peak over threshold approach under EVT. The returns are first filtered using GARCH-type models before the EVT analysis. The second stage of the IFM approach is the dependence modeling; this is done using D-vine copula. The D-vine copula is a class of regular vine copula that uses various bivariate copulas as building blocks. Empirical evidence using D-Vine copula for the dependence modeling shows that, both positive and negative dependence exist between pairs of tails, for both conditional and unconditional pairs. Some dependence parameters between upper and lower tails of two different assets in the portfolio are positive, showing that an extreme gain in one asset is associated with an extreme loss in another asset in the same portfolio.

keywords: D-Vine, Extreme Value, Pair-Copula, Dependence, GARCH.

*Corresponding author:reginasesay@yahoo.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

1 Introduction

Modeling dependence has always been a vital concept in financial management, especially financial risk management. Extreme occurrences such as extreme gains and extreme losses in financial market are unavoidable. Awareness of dependence in these extremes is a major step towards proper allocation of assets to a portfolio. Knowledge of dependence can also help investors to avoid possible losses by adjusting their portfolio mix accordingly.

Moreover, according to Kahneman and Tversky (1979), in their sturdy of human decision making, the pain people feel from a loss is about twice as strong as the pleasure felt from an equivalent experience of gain. As a result of this human behavior, when people are faced with investment decisions, they tend to have a stronger preference to avoid possible losses than for making gains and are willing to give up more to avoid possible losses. Investor are most likely to choose an outcome that has less risk of an experience loss and a lower expected return than an alternative choice that had greater risk of potential loss and a higher expected return.

In addition, it is well known that financial returns appear to have heavy tailed or leptokurtic distribution. Compared to the Normal or Gaussian distribution, return distributions tend to exhibit excess kurtosis (i.e. kurtosis greater than 3) indicating that returns have more mass in the tails than predicted by the normal distribution (MacKinlay et al., 1997) this paper, the extreme value theory is employed to model the tails of each asset return using peak over threshold (POT) approach. All sets of observations above (for upper tail) and below (for lower tail) the selected threshold for each asset are used in the dependence modeling using D-Vine Copula

Again, due to the difficulties associated with modeling dependence in multivariate return series, most of the financial literature makes use of multivariate normal and t distributions, despite their short comings. Recently, copulas have emerged as a powerful and flexible tool to create more flexible and realistic multivariate distribution in finance (Cherubini et al., 2004). However, although there exist a huge number of parametric bivariate copulas and also fairly large number of non-parametric bivariate copula, the range of higher dimensional copula is rather limited. Apart from the Gaussian and t copulas, one occasionally finds multivariate elliptical and Archimedean copulas. A number of empirical papers had shown that Gaussian copula do not fit financial returns data well (Mashal and Zeevi, 2002; Dobrić and Schmid, 2005). Both the Gaussian and the t copulas do not allow for different dependence structures between pairs of variables. A disadvantage of the student-t copula is it non-unimodality that makes it implausible for modeling joint returns (Schmidt and Theodorescu, 2006). Archimedean copulas seem to fit bivariate returns distribution well (Savu and Trede, 2008), but they are extremely restrictive in higher dimensional case, as they imply exchangeability and hence equicorrelated returns.

With the above difficulties in multivariate dependence modeling, this work used the D-vine copula to model the dependence structure in the multivariate extremes series. Vine copulas in general do not suffer from any of the above mentioned problems in the multivariate case, as it benefits from the wide variety of bivariate copulas use as building blocks. Vine copula was initially proposed by Joe (1996) and developed in more details in Bedford and Cooke (2001), Bedford and Cooke (2002) and in Kurowicka and Cooke (2006). Vines are flexible graphical models for describing multivariate copulas. They are build up using cascade of bivariate copulas (so-called pair copula), where each pair-copula can be chosen independently from the others. In particular, asymmetries and tail dependence can be taken in to account as well as (conditional) independence to build more parsimonious models. Their statistical breakthrough was due to the works of Aas et al. (2009) who described statistical inference technique for the two classes of canonical (C-) and D-vines.

2 Materials (Theoretical Framework)

This section contains theories and concepts used in the empirical analysis.

2.1 An overview of Extreme Value Theory (EVT)

Extreme value theory (EVT) is a branch of statistics that deals with extreme deviations from the median of probability distribution. It deals with events that have at least three things in common, they occur rarely, they are extreme in scope and they are difficult to predict. The EVT seeks to asses from a given ordered sample of a given random variable, the probability of events that are more extreme than any observed prior. In particular, it provides a theoretical framework for analyzing rare events and it has been known as the cousin of the well-known central limit theorem, as both theorems tell us what the limiting distributions are as the sample size increases. There are two main approaches for modeling extreme values, the block maxima approach and the peak over threshold (POT) approach (Embrocates et al., 1997). The second approach, the peak over threshold (POT) approach is used in the empirical analysis. We therefore discuss only the peak over threshold approach.

2.1.1 Peak over Threshold (POT) Approach

The POT approach consists of choosing a given threshold (high enough) and considering the extreme observations exceeding this threshold. The choice of the threshold is subject to a trade-off between the variance and bias. By increasing the number of observation in the tails, some observation from the center of the distribution are introduced in to the series and the index of the tail is more precise (less variance) but biased. On the other hand choosing a high threshold (for the upper tail) reduces the bias but makes the estimator more volatile (fewer observations).

Balkema and de Hann-Pickands Theorem: It is possible to find a positive measurable function β, where β is a function of u, such that:

$$\lim u \to_f (x) \sup_{(0 \leq x \geq x)} \left(|F_u(x) - G_{(\xi,\beta)}(x)| \right) \tag{1}$$

if and only if $F\epsilon$MDA$(H_\xi(x))$. That is, for a large class of underlying distribution F_u, as the threshold u gradually increases, the excess distribution function EQA converges to a generalized Pareto distribution.

2.1.2 The Generalized Pareto distribution

The Generalized Pareto distribution (GPD) is the limiting distribution of the peak over threshold approach and is defined as:

$$G_{\xi,u,\beta}(x) = \begin{cases} 1-(1+\xi\frac{x-u}{\beta})^{\frac{-1}{\xi}} & \text{if } \xi \neq 0 \\ -e^{(-\frac{x-u}{\beta})} & \text{if } \xi = 0 \end{cases}$$

with

$$x = \begin{cases} u, \infty & \text{if } \xi \geq 0 \\ u, u-\frac{\beta}{\xi} & \text{if } \xi < 0 \end{cases}$$

where ξ is the shape parameter. β is the scale parameter, and u is the location parameter. Just like the Generalized Extreme Value distribution, the GPD distribution contains three distributions as particular cases: Ordinary Pareto distribution: $\xi = \alpha^{-1} > 0$, Pareto type distribution $\xi = 0$ Exponential distribution: $\xi = \alpha^{-1} < 0$

2.2 Copula Theory

Copula, according to Nelson (2006), can be explained from two points of view, first, copulas are functions that join or couple multivariate distribution functions to their one dimensional marginal distribution function. Second, Copulas are multivariate distribution functions whose dimensional margins are uniform on the interval [0,1]. Practically, a copula is often used to construct joint distribution function by combining the marginal distributions and the dependence between the variables.

Properties: We only focus on bivariate copulas, as bivariate copulas are used as building blocks in modeling dependence using D-vine copula. A 2-dimensional copula is a function c: $[0,1]^2 \rightarrow [0,1]$ with the following properties:

1. For every$u \in [0,1], C[0,u] = C(u,0) = 0$
2. For every $u \in [0,1]$, $C(u,1) = u$ and $C(1,u) = u$
3. For every (u_1,u_2),(v_1,v_2) $\in [0,1][0,1]$ with $u_1 \leq v_1$ and $u_2 \leq v_2$: $C(v_1,v_2) - C(v_1,u_2) - C(u_1,v_2) + C(u_1,u_2)0$

Property 1 is referred to as the grounded property of copula. It says that the joint probability of both outcomes is zero if the marginal probability of any out comes is zero; Property 3 is the two-dimensional analogue of a non-decreasing one-dimensional function. A function with this feature is therefore called 2-increasing.

Sklar's Theorem: Let H be a 2-dimentional joint distribution function with marginal distributions F and G. There exists a copula C such that:

$$H(x, y) = C(F(x), G(y))$$

, If F and G are continuous, then C is unique. Conversely, if C is a copula and F and G are distribution functions, then the function H is a joint distribution function with marginal distributions F and G. The converse of this theorem implies that the combination of two different marginal distributions with any copula will define a valid bivariate distribution.

Corollary: Let H be a 2-dimentional joint distribution with continuous marginal distribution F and G with copula C, Satisfying Sklar's theorem, then for

$$u \in [0, 1] \ \& \ v \in [0, 1]$$

and

$$v \in [0, 1] : C(u, v) = H(F^{(-1)}(u), G^{(-1)}(v))$$

where F^{-1} and G^{-1} denote the inverses of F and G. This corollary provides a motivation for calling a copula a dependence structure as it links the quantiles of two distributions rather than the original variables.

2.2.1 Conditional Copula

The conditional copula of $(X, Y \mid W) = w$ where

$$X \mid W = w \sim F_1(X|W)(. \mid w) \text{ and } Y \mid W = w \sim F_1(Y \mid W)(. \mid w)$$

is the conditional joint distribution function of $U \approx F_1(X \mid W)(X \mid w)$ and

$$V \approx F_1(Y \mid W)(Y \mid w) \text{ given } W = w.$$

The two variables U and V are known as the conditional "probability integral transforms"of X and Y given W. These variables have Uniform (0,1) distributions, regardless of the original distribution of X and Y (Fisher, 1925; Rosenblatt, 1952). Patton (2002) shows that a conditional copula has all of the properties of the unconditional copula.

As bivariate copulas are used as building blocks in the dependence modeling using D-Vine, we next present the Fréchet-Hoeffding bounds for bivariate copulas.

2.2.2 Fréchet Hoeffding copula bounds

Due to the existence of some extreme cases of dependence, Frèchet and Hoeffding independently showed that, a copula always lies in between certain bounds.

Theorem: For any sub-copula $C^{'}$ with domain $S_1 \times S_2$ and for any $(u, v) \in S_1 \times S_2$

$$C^{-}(u,v) := \max(u+v-1,0) \leq C^{'}(u,v) \leq min(u,v) =: C^{+}C^{-}$$

and

$$C^{+}$$

are the Fréchet-Hoeffding lower and upper bound respectively. The Frèchet Hoeffding lower bound corresponds to perfect negative dependence and the upper bound corresponds to perfect positive dependence. Random variables say,$X_1, \ldots, X_d$ are said to be comonotonic if their copula corresponds to C^{+} and countermonotonic if their copula corresponds to C^{-}. Some copulas, like the Clayton copula attains only the Frechet upper bounds (as $\theta \to \infty$) while others like the Frank Copula attain both the Frechet's upper and lower bounds(as $\theta \to \infty$ and $\theta \to -\infty$ respectively).

2.2.3 Monte Carlo integration for copula models

Consider the random vector $(X_1, \ldots, X_d)$ with cdf H, the expectation of a response function $g : R^d \to R$ applied to this random vector can thus be written as

$$Eg(X_1, \ldots, X_d) = \int {}^{d}_{R} g(x_1), \ldots, x_d) dH(x_1, \ldots, x_d)$$

If H is given by a copula model and the copula, C is absolutely continuous, i.e. C has a density c, this expectation can be rewritten as

$$E[g(X_1, \ldots, X_d)] = \int_{[0,1]^d} g\left(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\right)) c(u_1, \ldots, u_d) du_1, \ldots, du_d$$

If copula and margins are known, the expectation can be approximated through the following Monte Carlo algorithm:

1. Draw a sample $(U_1^k, \ldots, U_d^k) \sim C(K = 1, \ldots, n)$ of size, n from the copula c.
2. Produce a sample of $(X_1, \ldots, X_d)$ by setting $(X_1^k, \ldots, X_d^k) = (F_1^{-1}(U_1^k), \ldots, F_d^{-1}(U_d^k) \sim H(k = 1, \ldots, n)$
3. Approximate $E[g(X_1, \ldots, X_d)]$ by its empirical value: $E[g(X_1, \ldots, X_d)] \approx \frac{1}{n} \sum_{i=1}^{n} g(X_1^k, \ldots, X_d^k)$

2.2.4 Pair-Copula Decomposition of a multivariate Distribution

Pair copulas, originally introduce by Joe (1996) provide a flexible way to construct multivariate distribution. In using vine copula, the first step is to decompose the multivariate copula into a cascade of bivariate copulas using the idea of pair copula decomposition. Hence the multivariate density with dimension d can be factorized as

$$f(x_1,\ldots,x_d) = f(x_d|x_1,\ldots,x_{d-1}).f(x_1,\ldots,x_{d-1}) = f(x_d).f(x_{d-1}|x_d)\cdot$$
$$\cdot f(x_{d-2}|x_{d-1},x_d),\ldots,f(x_1 \mid x_2,\ldots,x_d)$$

To get a pair copula for $f(x_1 \ldots x_d)$, we only need to replace each of the conditional densities bit by bit with products of pair-copulas and marginal densities. The following examples illustrate the construction of pair copula. For d=3, the joint 3-dimentional density can be decompose as:

$$f(x_1,x_2,x_3) = f_1(x_1).f_{2|1}(x_2|x_1).f_{3|12}(x_3|x_1,x_2)$$

In general,

$$f(x_{d-2}|x_{d-1},x_d),\ldots,f(x_1 \mid x_2,\ldots,x_d)f(x_j \mid x_i = \frac{f(x_i,x_j)}{f_i(x_i)} = f(x_i,x_j) \mid (f_i(x_i)) =$$
$$= c_ij(F_i(x_i),F_j(x_j)).f_j(x_j).f_i(x_i))$$

The multivariate density factorization can thus be express in terms of the marginal densities and a product of $\frac{d(d-1)}{2}$ (d is the dimension of the multivariate density) bivariate pair-copulas, using the following formula iteratively:

$$f(x \mid v) = c_{xv_j|v_j}|v_{-j}\left(F_{x|v_{-j}}(x \mid v_{-j}), F_{v_j|v_{-j}}(v_j \mid v_{-j}\right).f(x \mid v_{-j})$$

where $v = (v_1,\ldots,v_d)$ is a vector v_j with $j = 1,\ldots,d$ the j^{th} element of this vector and $v_{-j} = (v_1 \ldots v_{j-1}, v_{j+1} \ldots, v_d)$ the vector v without the $j-th$ component. Hence, a four dimensional density ($d = 4$) can be decomposed into $\frac{4(4-1)}{2} = 6$ pair- copulas and four marginal densities.

The conditional marginal distributions, $F(x \mid v)$ can be calculated reclusively. Joe (1996) showed that, for every j,

$$F(x \mid v) = \frac{\delta c_{x,v_j|v_{(-j)}}(F(x| \mid v_{-j}).F(v_j \mid v_{-j}))}{\delta F(v_j|v_{-j})}$$

where $C_(ij|k)$ is a bivariate copula distribution function. When v is univariate and x and v are uniformly distributed on the [0,1] interval, we have

$$F(x \mid v) = \frac{\delta C_{xv}(x,v)}{\delta v}$$

The function $h(x,v,\Theta)$ is called the $h(.)$ function, and is:

$$h(x,v,\Theta) = F(x \mid v) = \frac{\delta C_{xv}(x,v)}{\delta v}$$

Here, Θ represent the set of parameters for the copula of the joint distribution function of x and v. The inverse of this h(.) function, $h^{-1}(x,v,\Theta)$ with respect to the first variable x, is needed for sampling from pair-copula distribution.

2.2.5 Vine Structure

Vines are graphical representations to specify pair copula construction. The C- and D-vines are most appropriate if their structure is explicitly motivated by the data. There are many ways to decompose multivariate density function in to pair copula. However Bedford and Cooke (2001) and Bedford and Cooke (2002) have introduced a graphical model called a regular vine to help organize all possible decompositions.

The D- vines are constructed by choosing a specific order of the variables. For example, in the first tree, the dependence of the first and second; second and third; of the third and fourth, and so on are modeled using pair-copula. I.e., if we assume the order 1, . . . , d we can model the pair $(1, 2), (2, 3), (3, 4)$,etc. In the second tree, conditional dependencies of the first and third given the second variable (the pair $(1, 3 \mid 2)$), the second and fourth given the third (the pair $(2, 4 \mid 3)$), and so on are modeled. The pair dependencies of variables are modeled in the same way in subsequent trees conditioned on those variables which lie between the variables in the first tree, e.g. the. pair $(1, 5 \mid 2, 3, 4)$.

Generally, for regular vine with dimension d, there are d-1 pairs at level 1 (they satisfy the condition that the graph formed with edges based on the pairs has no cycle), $d - 2$ pairs in level 2 etc. and each pair has at least one element in common. For $k = 2, \ldots, d - 1$, there are d-k pairs in level k. For other conditions of regular vine tree construction, see Bedford and Cooke (2001); Bedford and Cooke (2002).

3 Methodology

3.1 Copula Estimation

There are two popular parametric estimation methods available for copula estimation, the full maximum likelihood (FML) and the inference function for margin (IFM). The first method simultaneously estimates all the parameters of the marginal distributions and the copula using full maximum likelihood (FML). However in some situation, the maximum likelihood estimation may be difficult to obtain as both the marginal and copula parameters must be estimated jointly. Therefore, numerical methods have to be adopted to solve the optimization problem, which slows down computation and often leading to convergence difficulties. Due to the problem face by using the FML in the curse of estimation, this paper uses the inference function for margins (IFM) method in the dependence modeling using D-vine copula.

The IFM is a two step estimation procedure and involves the maximum likelihood estimation of the dependence parameter given the estimated marginal distributions. In the first step, the parameters of the marginal distributions are estimated. In the second step the copula parameters are estimated conditioned on the previous marginal distributions estimates. This method exploits an attractive feature of copula for which the dependence structure is independent of the marginal distributions. Under regularity conditions, Patton (2006) showed that the IFM estimator is consistent and asymptotically Normal. Below is an illustrative description of the IFM approach.

We know that from Sklars Theorem, the density f of d-dimensional F with univariate margins $f_1, f_2, \ldots, f_d$ can be represented as

$$f(x_1, x_2, \ldots, x_d) = c(F_1(x_1), F_2(x_2), \ldots, F_d(x_d) \prod_{i=1}^{d} f_i(x_i)$$

where

$$c(u_1, u_2, \ldots, u_d) = \frac{(\delta C(u_1, u_2, \ldots, u_d))}{(\delta u_1 \delta u_2, \ldots, \delta u_d)}$$

is the density of the d-dimensional copula $C(u_1, u_2, \ldots, u_d;)$. This implies the following decomposition of the log-likelihood function, $L = \sum_{j=1}^{n} \log f\left(x_1^{(j)}, x_2^{(j)}, \ldots, x_d^{(j)}\right)$ of a random sample of (i.i.d.) vector $x^j = (x_1^{(j)}, x_2^{(j)}, \ldots, x_d^{(j)}), j = 1, 2, \ldots, n$, with the density f.

$$L = \binom{L_c}{dependence} + \binom{\sum_{i=1}^{d} L_i}{marginals}$$

This decomposition reflects the IFM approach used in this paper. From the above decomposition:

$$L_c = \sum_{j=1}^{n} \log\left(c(F_1(x_1^{(j)}, F_2(x_2^{(j)}, \ldots, F_d({x_d}^{(j))}\right)$$

$$L_i = \sum_{(j=1)}^{n} logf_i\left(x_i^{(j)}\right), i = 1, 2, \ldots, d$$

L_c is the log-likelihood contribution from the dependence structure of the data represented by the copula C whereas, L_i is the log-likelihood contribution in L from each margin.

In modeling the marginal distributions, under stage one of the IFM, we get the estimators

$$\hat{\alpha_i}^{IFM} = argmax_{\alpha_i} L_i(\alpha_i).$$

Stage two of the IFM approach is the dependence modeling, where the estimator $\hat{\theta}^{IFM}$ of the copula parameter $\hat{\theta}^{IFM}$ is computed by maximizing the copula likelihood contribution L_C, with the marginal parameters α_i replaced by their first-stage estimators i.e.:

$$\hat{\alpha_i}^{IFM} : \hat{\theta}^{IFM} = argmax_{\theta} L_C(\alpha_1^{IFM}, \hat{\alpha_2}^{IFM}, \ldots, \alpha_d^{IFM}, \theta)$$

The two-stage IFM estimator $\left(\hat{\alpha_1}^{IFM}, \hat{\alpha_2}^{IFM}, \ldots, \hat{\alpha_d}^{IFM}, \hat{\theta}^{IFM}\right)$ solves:

$$\left(\frac{\partial L_1}{\partial \alpha_1}, \frac{\partial L_2}{\partial \alpha_2}, \ldots, \frac{\partial L_d}{\partial \alpha_d}, \frac{\partial L}{\partial \theta}\right) = 0$$

3.2 Modeling of Marginal Distributions

The first stage of the IFM approach discussed above is the modeling of the Marginal distribution. In this paper, the peak over threshold approach is used to model each marginal distribution. Since EVT assumes independent and identically distributed observations, we first filter each return series using GARCH-type models before the extreme value analysis.

3.2.1 Fitting excess over a threshold

In fitting the excess over a threshold, the first step is to choose the threshold. In this paper, we used the mean residual life plot to choose the threshold. The threshold is chosen where the plot is approximately linear. For this, let u be the chosen threshold, $X_1 \dots X_n$ the random variables exceeding this threshold with a distribution $F \in MDA(H_\xi))$, and $Y_{u_1} \dots Y_{u_i}$ the series of exceedances (where $Y_i = X_i - u$). The distribution of excess beyond u is then given by:

$$F_u(y) = p(X - u \leq y \mid x > u) = p(Y \geq y \mid X > u),\ y \leq 0$$

and the distribution, F of the extreme observations, X_i is given by

$$F(u + y) = p(X \geq u + y) = p(X \geq u + y \mid X > u) \cdot p(X > u)$$
$$F(u + y) = p(X - u \leq y \mid X > u) \cdot p(X > u)$$
$$F(u + y) = F_u(y) \cdot F(u)$$

This allows us to estimate the tail of the original distribution by estimating F and F_u seperately. According to Pickands III (1975), Balkema and De Haan (1974), for a high threshold u,

$$\bar{F}_u(y) = \bar{G}_{\xi\beta(u)}(y)$$

$F(u)$ can be estimated from the empirical distribution of the observations:

$$\hat{F}_u = \frac{1}{n}\sum_{i=1}^{n} I_{X(i>u)} = \frac{N_u}{n}$$

Therefore

$$\hat{F}(u + y) = \frac{N_u}{n}\left(1 + \hat{\xi}\frac{y}{\beta}\right)^{\frac{-1}{\xi}}$$

Next, we estimate the parameters, ξ and β using MLE.

3.3 Dependence Modeling

The second stage of the IFM approach is the dependence modeling. This is done using D-vine copula.

3.3.1 Parameter Estimation of D-Vine Copula

Classically, parameters of statistical models are often estimated using ML techniques, but joint maximum likelihood estimation of the D-vine copula parameter is very challenging, as the vine decomposition involves $\frac{d(d-1)}{2}$ bivariate copulas. However, this problem can be tackled by sequential estimation. In this paper, we adopted the algorithm presented in Aas et al. (2009) for the numerical maximization of the log-likelihood with adequate starting values. The starting values are obtained using sequential estimation approach. For the D-vine copula, we used the following steps in the sequential estimation:

1. First, we reorder the variables (data sets) such that variables that are more correlated are place side by side.
2. Next, we determined which pair copulas to use in the first tree. The copulas are selected using AIC. Base on the direction of dependence (positive or negative) between the variables as indicated by Kendalls tau values, we select candidate copulas. The criteria are computed for all candidate copula families and the family with the minimum AIC value is chosen for each pair.
3. We estimated the copula parameters of the first tree using uniformly transformed data. All pair-copula parameters are estimated using MLE technique.
4. We then used the parameter estimates from the first tree and the appropriate h (.) function to compute pseudo observations for the second tree.
5. Next, we used AIC to determine which pair copulas to use in the second tree and then estimate the pair copula parameters.
6. Compute the implied observation for the third tree using the estimated copula parameters from the second tree and the appropriate h(.) function.
7. This process continued to the last tree.

3.3.2 Parameter Estimation via Maximum Likelihood

The sequential estimates are used as starting values in the numerical maximization of the log-likelihood The D-vine copula log-likelihood with parameter set $_DV$ is:

$$l_{DV}(\theta_{DV|u}) = \sum_{k=1}^{n}\sum_{i=1}^{d-l}\sum_{j=1}^{d-j} log[c_{j,j+i|(j+1),\ldots,(j+i-1)}(F_{j|j+1,\ldots\ldots(j+i-1)}, F_{(j+i)|(j+1),\ldots,(j+i-1)} \mid$$

$$\mid \theta_{(j,j+i|(j+1),\ldots,(j+i-1)})$$

We next present the parameter ranges of some copulas used in the empirical analysis, together with their tail dependencies and Kendall tau values.

Table 1: Functions, ranges of dependent parameters, corresponding Kendall tau values and Tail dependence

Copula	Function C(u,v)	Parameter Range θ
Gumbel Copula	$\exp\left(-((-\ln(u))^{\theta}+(-\ln(v))^{\theta})^{\frac{1}{\theta}}\right)$	$[1,\infty)$
Frank Copula	$-\frac{1}{\theta}\ln\left(1+(\frac{e^{\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$	$(-\infty,\infty)\mid\{0\}$
Joe Copula	$1-((1-u)^{\theta}+(1-v)^{\theta}-(1-v)^{\theta})^{\frac{1}{\theta}}$	$\theta>1$
Clayton Copula	$\left(u^{-\theta}+v^{-\theta}-1\right)^{\frac{-1}{\theta}}$	$\theta>0$
Gaussian Copula	$\Phi_G[\Phi^{-1}(u),\Phi^{-1}(v);\theta]$	$-1<\theta<1$

Copula	Kendalls tau	Tail Dependence (Lower,Upper)
Gumbel Copula	$1-\frac{1}{\theta}$	$(0,2-2^{\frac{1}{\theta}})$
Frank Copula	$1-\frac{4}{\theta}+4\frac{D_i(\theta)}{\theta}$	$(0,0)$
Joe Copula	$1+\frac{4}{\theta^2}\int_0^1 tlog(t)(1-t)^{2(1-\theta)/\theta}dt$	$(0,2-2^{\frac{1}{\theta}})$
Clayton Copula	$\frac{\theta}{\theta+2}$	$(2^{-\frac{1}{\theta}},0)$
Gaussian Copula	$\frac{2}{\pi}\arcsin(\theta)$	(0,0)

In addition, we also used some rotated versions of Gumbel, Clayton and Joe copulas. Rotation of 90 degrees and 270 degrees allows for modeling negative dependence, which is not possible with the non-rotated versions. Table 2 contains the parameter ranges of these rotated copulas.

Table 2: Copulas and their parameter Ranges

Rotated Copula	Parameter Range
Rotated Clayton (90 and 270 degrees)	$(-\infty,0)$
Rotated Gumbel (90 and 270 degrees)	$(-\infty,-1)$
Rotated Joe (90 and 270 degrees)	$(-\infty,-1)$

4 Application (Empirical Analysis)

4.1 Description of Data and Preliminary Tests

To model the extreme dependence in the tails, this work considers a portfolio consisting of three assets, which are: the Nikkei 225 (N225) more commonly called the Nikkei Stock Average which is a stock market index for the Tokyo Stock Exchange and the most widely quoted average of Japanese equities; the French stock market index known as Cotation Assistée en Continu (CAC40) which is one of the main national indices of the pan-European stock exchange and the German Stock Index which was formally known as Deutscher Aktien index (DAX). Each series consist of the daily closing prices of assets from 1990/3/16 to 2014/4/2, with 5913 observations in each series. The daily returns, are calculated as:

$$r_t = log\frac{P_t}{P_{t-1}} \times 100$$

where r_t is the return at time t and P_t the price at time t.

4.1.1 Histogram Normality Test and Descriptive Statistics

The normality tests and descriptive statistics for the three series are presented in figure 2. Like most financial time series, all the three series exhibit evidence of fat tails, with the kurtosis of each series greater than 3. The Jarque-Bera normality test further confirms that all the three series are not normally distributed.

4.1.2 Exploratory Data Analysis (Q-Q plot)

The Quantile-Quantile (Q-Q) plot is a graphical method for comparing two probability distributions by plotting their quantiles against each other. A concave departure from the straight line in the Q-Q plot is an indication of a heavytailed distribution, whereas a convex departure is an indication of a thin tail. To be precise, when comparing the data (residuals) to the normal distribution, if the plot curve to the right, the data have right tail which is heavier than the normal distribution and if it curves up to the left, the data have a left tail that is heavier than the normal distribution. Figure 3 shows that in all the three series, the standardized residuals do not follow the normal distribution. The distribution of each residual has a tail that is heavier than that of the normal distribution.

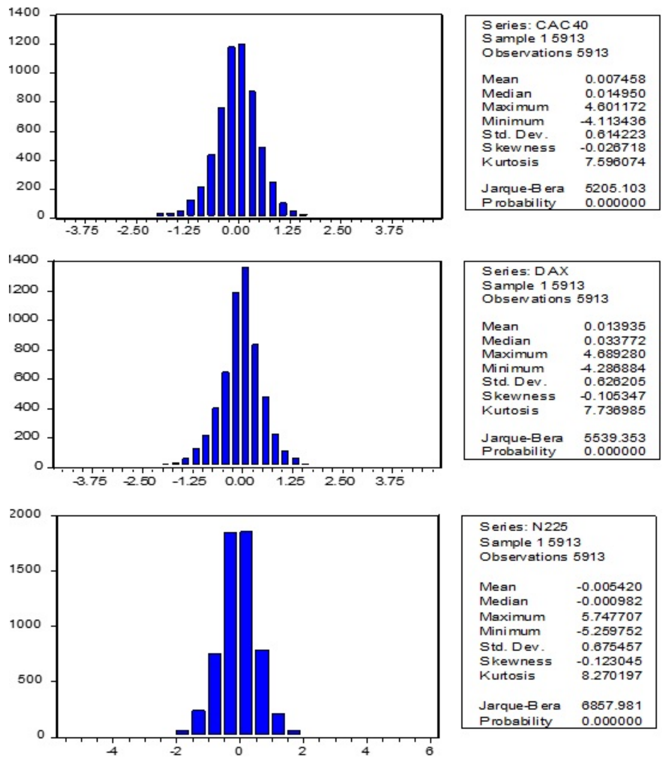

Figure 1: Histogram Normality Test

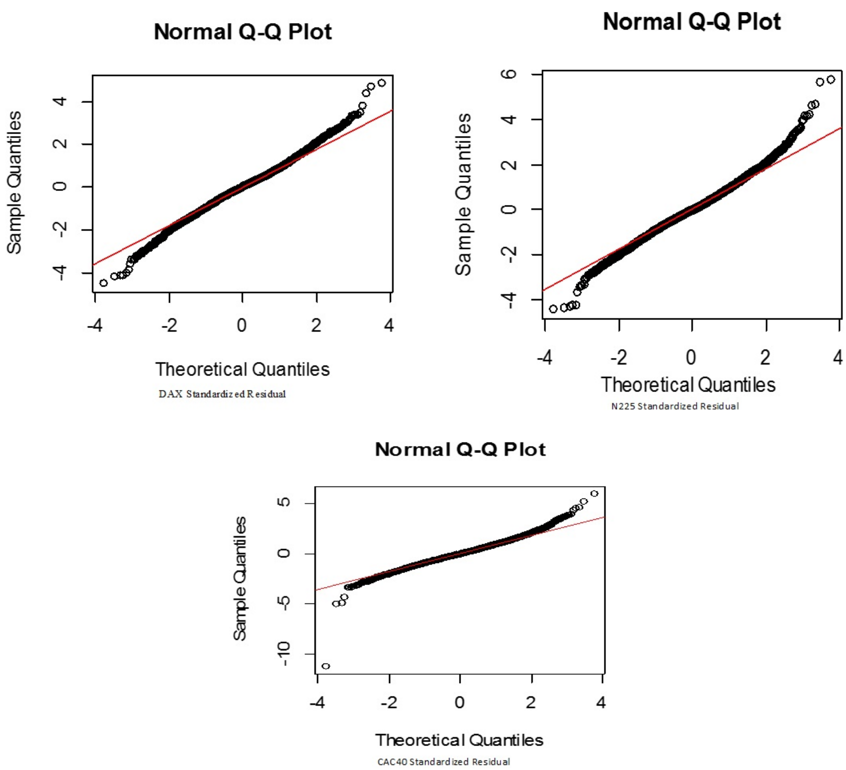

Figure 2: QQ-Plots of Residuals against the Normal Distribution

4.2 ARCH LM Test

This test is a Lagrange multiplier test for autoregressive conditional heteroskedasticity (ARCH) in the residuals (Engle, 1982). The obs*R-squared statistic is Engles LM test statistic and is computed as the number of observation times the R2 from the test regression. The LM statistic is asymptotically distributed with 2 (q) under general condition. Under the LM test, the null hypothesis of no ARCH up to order q is rejected if TR^2 is greater than the $^{2}_{q}$ tabulated value at a given significance level.

Table 3: (G) ARCH Effect (Engels LM) Test on Residuals

Asset	N225	CAC40	DAX
Selected GARCH(p,q) model	GARCH(2,1)	GARCH(1,2)	GARCH(2,1)
TR^2	0.087554	0.239079	0.157572
p-value	0.767310	0.624872	0.691402

The selected GARCH models along with the ARCH LM test statistics on the residuals for the three series are presented in Table 3. The ARCH LM test confirms that the ARCH effect has been mitigated at the 1% significance level. The residuals are now approximately iid and therefore can be used for extreme value Analysis.

4.3 Split Sample

To explore the asymmetries normally found in financial returns, we split each residual series into two (Lower and Upper tail). To achieve this, we selected about 25% of the observations from the tail ends (upper and Lower) of the distributions of each series. Samples selected from the upper tail end of the distributions are referred to as upper tails and those selected from the lower tail ends are referred to as lower tails. This is just a split sample, the actual determination of extreme observations for each of these split samples is done using the peak over threshold approach under extreme value theory. We call the split samples lower and upper tails because, the observations included in each sample are selected from the tails of each distribution. The signs of the residuals for the lower tails in each series are change to positive values. One thousand four hundred and seventy eight (1478) observations are included in each split sample. These split samples are then used in the empirical analysis.

4.4 Threshold Determination

Threshold Determination: We used the mean residual life plots to determine the threshold. The mean residual life plot involves plotting the threshold (u) against the mean excess for range of values of the threshold. The plot should be linear above the threshold at which the GPD model becomes valid. McNeil et al. (2015) show that for a sample size of 1,000 fixing the number of threshold exceedance to 100 yields good estimate of

Value-at-Risk (VaR) and Expected Shortfall (ES). We therefore, choose the threshold such that about 10% of the observations are included in each tail. Figure 3 shows the mean residual life plots for each tail.

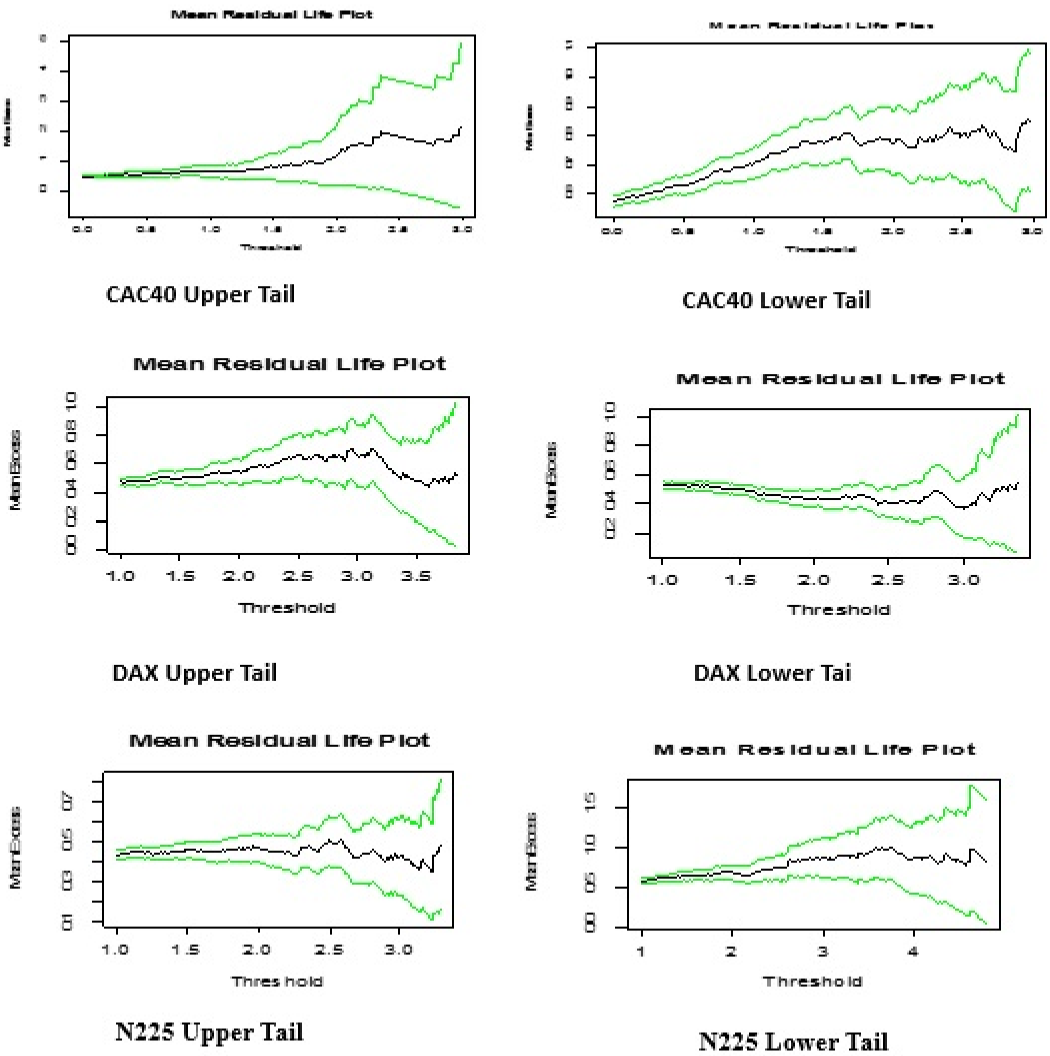

Figure 3: Mean Residual life Plots

Table 4 shows the selected threshold values together with the number of exceedances for each threshold and their corresponding sample quantiles. The MLE estimates of the shape and scale parameters together with their standard errors are also presented in this table. We can see that apart from N225 lower tail with a shape parameter of 0.05249 which is not significantly different from zero, all the other shape parameters are greater than zero ($i.e. \xi > 0$). Precisely, all the shape parameters are positive which means that the GPD is equivalent to the Pareto distribution.

Table 4: Maximum Likelihood Parameter Estimation

UPPER TAIL	CAC40	DAX	N225
No. of Obs.	1478	1478	1478
U	0.54	1.4	1.2
Proportion Of Exceedances—	0.1084	0.1057	0.1023
Scale (β) (Se)	0.3884 0.04327	0.4602 0.04864	0.2448 0.03511
Shape (ξ) (se)	0.2081 0.08022	0.1314 0.06980	0.4572 0.12380
LOWER TAIL	CAC40	DAX	N225
No. of Obs.	1478	1478	1478
U	0.1016	0.9	1.22
Exceedances	150	152	159
Proportion of Exceedances	0.1016	0.103	0.108
Scale (β) (se)	0.3453 0.04914	0.4279 0.04598	0.59899 0.0836
Shape(ξ) (se)	0.3056 0.11920	0.1744 0.07187	0.05249 0.1012

Table 5: D-Vine Tree structure with Best Copula Fits, Parameter Estimates and Standard Errors

Tree1	u_{DL},u_{CU} Fam=34 Par= -1.9898071 S.E= 0.008397435	u_{CU}, u_{DU} Fam=4 Par= 1.876682 S.E=0.006855861	u_DU, u_CL Fam=5 Par= 0.49112032 S.E= 0.01454904	u_{CL}, u_{NU} Fam=5 Par=0.3825851 S.E=0.0830336	u_NU, u_NL Fam=34 Par= -1.8979801 S.E= 0.0085
Tree 2	$u_{DL}, u_{DU} \vert u_{CU}$ Fam=26 Par=-1.891191 S.E= 0.02735181	$u_{CU}, u_{CL} \mid u_{DU}$ Fam=3 Par= 0.9607828 S.E=0.02721778	$u_{DU}, u_{NU} \mid u_{CL}$ Fam=5 Par= 0.457956 S.E= 0.03286678	$u_{CL}, u_{NL} \mid u_{NU}$ Fam=5 Par=0.3964029 S.E=0.0853806	
Tree 3	$u_{DL}, u_{CL} \mid u_{CU}, u_{DU}$ $Fam = 36$ $Par = -1.490849$ $S.E = 0.02081305$	$u_{CU}, u_{NU} \mid u_(DU,)u_{CL}$ $Fam = 13$ $Par = 0.398203$ $S.E = 0.02145409$	$u_{DU}, u_{NL} \mid u_{CL} u_{NU}$ $Fam = 13$ $Par = 0.06016624$ $S.E = 0.01435643$		
Tree 4	$u_{DL}, u_{NU} \mid u_{CU}, u_{DU}, u_{CL}$ $Fam = 26$ $Par = -1.12463$ $S.E = 0.009023852$	$u_{CU} u_{NL} \mid u_{DU}, u_{CL}, u_{NU}$ $Fam = 34$ $Par = -1.000145$ $S.E = 0.007236771$			
Tree 5	$u_{DL}, u_{NL} \mid u_{CU}, u_{DU}, u_{CL}, u_{NU}$ $Fam = 3$ $S.E = 0.002139803$				

where u_{DL} = DAX Lower Tail, u_{NU} = N225 Upper Tail, u_{CL} = CAC 40 Lower Tail, u_{DU} = DAX Upper Tail, u_{NL} = N225 Lower Tail and Fam5 = Frank Copula, Fam26 = Rotated Joe Copula (90 degrees), Fam4 = Gumbel Copula, Fam13 = Rotated Clayton Copula (180 degrees: survival Clayton), Fam34 = Rotated Gumbel Copula (270 degrees), Fam3 = Clayton copula, Fam36 = rotated Joe copula (270 degrees).

5 Results and Discussion

From Table 5, we can see that both positive and negative dependence exist between the pairs of tails. Positive dependence between two tails, say lower tail (losses) and upper tail (gain) implies that an increase in the losses of one asset is associated with an equivalent increase in the gain of the other asset. As a result, the contribution of the two assets to the expected value of the portfolio is negligible. For example in tree 1, the dependence between DAX upper tail and CAC40 lower tail is positive (Par= 0.49112032). This dependence parameter is estimated using Fam 5, which is the Frank copula. This copula takes values from $-\infty$ to ∞ (more precisely,$(-\infty, \infty) \mid \{0\}$). A zero value implies independence. Similarly in this same tree, the dependence between CAC40 upper tail (u_{CU}) and DAX upper tail (u_{DU}) is positive ($i.e., par = 1.876682$).This particular dependence parameter is estimated using Fam4 which is the Gumbel copula. Gumbel copula takes values from 1 to infinity ($i.e., [1, \infty)$), whereas a value of 1 implies independence. When the estimated parameter is above one, the two tails are somehow dependent.

In trees 2, 3, 4 and 5, some of the conditional dependence parameters are estimated using rotated copulas. For example, in the second tree, conditioned on CAC40 upper tail (u_{CU}) the dependence between DAX lower tail (u_{DL}) and DAX upper tail (u_{DU}) is negative ($Par = -1.891191$). This dependence parameter is estimated using Fam26 which is the rotated Joe Copula (90 degrees) which takes vales from $-\infty$ to -1 . The more negative the dependence parameter, the greater the strength of the negative dependence. Also, in tree 3 conditioned on CAC40 upper tail (u_{CU}) and DAX upper tail (u_{DU}), the dependence between DAX lower tail (u_{DL}) and CAC40 lower tail (u_{CL}) which has been estimated using Fam36, that is the rotated Joe copula (270 degrees) turns out to be negative. This copula takes values from $-\infty$ to -1 ($i.e., \theta \epsilon (-\infty. -1)$). As the copula parameter, θ approaches ($-\infty$) the copula attains the Fre chet-Hoeffding lower bound, which corresponds to perfect negative dependence and as $\theta \to -1$ the two tails become independent. The estimated parameter (-1.490849) shows that the two tails are negatively dependent.

Finally, in the last tree, conditioned on CAC40 upper tail (u_{CU}, DAX upper tail (u_{DU}), CAC40 lower tail (u_{CL}) and N225 upper tail, the dependence between DAX lower tail (u_{DL}) and N225 lower tail (u_{NL}) estimated using Fam3, which is the Clayton copula, is positive,. This copula only measures positive dependence and takes values from 0 to ∞ (i.e.,$\theta \epsilon (0, \infty)$. As θ goes to zero, the two tails become independent. The estimated dependence parameter ($Par = 0.0001458298$) in tree 5 shows that with all the other assets in the portfolio, the dependence between DAX lower tail and N225 lower tail is very weak.

6 Conclusion

As the purpose of this work is to explore the dependence between the lower and upper tails of 3 assets in a portfolio using EVT and D-vine, the above empirical evidence shows that both positive and negative dependence exist between pairs of tails for both conditional and unconditional pairs. Bivariate copulas are used as building blocks in modeling the dependence using D-vine. Depending on the type (positive or negative) of dependence, different types of bivariate copulas were used to estimate the dependence parameters. Some estimated dependence parameters between upper and lower tails of two different assets in the portfolio are positive, showing that an extreme gain in one asset is associated with an extreme loss in another asset. The estimated dependence parameters further point out that there exist some forms of dependence between the lower and upper tails of the assets contained in the portfolio that worth taking into account for good management practice. Base on this investigation, our recommendation to managers and investors is to pay greater attention to the extreme dependence in the lower and upper tails (extreme losses and extreme gains) while mixing assets in the portfolio, to avoid the risk of extreme losses whiles chasing extreme gains.

References

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44(2):182–198.

Balkema, A. A. and De Haan, L. (1974). Residual life time at great age. *The Annals of probability*, pages 792–804.

Bedford, T. and Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence*, 32(1-4):245–268.

Bedford, T. and Cooke, R. M. (2002). Vines: A new graphical model for dependent random variables. *Annals of Statistics*, pages 1031–1068.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula methods in finance*. John Wiley & Sons.

Dobrić, J. and Schmid, F. (2005). Testing goodness of fit for parametric families of copulasapplication to financial data. *Communications in StatisticsSimulation and Computation®*, 34(4):1053–1068.

Embrocates, P., Kluppelberg, C., and Mikosch, T. (1997). Modelling extreme events for insurance and finance.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Fisher, R. A. (1925). *Statistical methods for research workers.* Genesis Publishing Pvt Ltd.

Joe, H. (1996). Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters. *Lecture Notes-Monograph Series*, pages 120–141.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291.

Kurowicka, D. and Cooke, R. M. (2006). *Uncertainty analysis with high dimensional dependence modelling.* John Wiley & Sons.

MacKinlay, A. C., Lo, A. W., and Kampbell, J. (1997). The econometrics of financial markets.

Mashal, R. and Zeevi, A. (2002). Beyond correlation: Extreme co-movements between financial assets. *Unpublished, Columbia University.*

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools.* Princeton university press.

Nelson, R. (2006). An introduction to copula in american mathematical society. *Providence, RI*, pages 271–294.

Patton, A. (2002). Skewness, asymmetric dependence. Technical report, and portfolios. mimeo.

Patton, A. J. (2006). Estimation of multivariate models for time series of possibly different lengths. *Journal of applied econometrics*, 21(2):147–173.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics.*

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The annals of mathematical statistics*, 23(3):470–472.

Savu, C. and Trede, M. (2008). Goodness-of-fit tests for parametric families of archimedean copulas. *Quantitative Finance*, 8(2):109–116.

Schmidt, R. and Theodorescu, R. (2006). On the schur unimodality of copulas and other multivariate distributions. In *Seminar of Economic and Social Statistics, University of Cologne, Working Paper.*